

## SEQUENCE LISTING

<110> Cox III, George Norbert  
Case, Casey Christopher  
Eisenberg, Stephen P.  
Jarvis, Eric Edward  
Spratt, Sharon Kaye  
Sangamo Biosciences, Inc.

<120> Regulation of Endogenous Gene Expression in Cells Using  
Zinc Finger Proteins

<130> 019496-002200US

<140> 09/229,037  
<141> 1999-01-12

<160> 40

<170> PatentIn Ver. 2.0

<210> 1  
<211> 25  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: exemplary motif  
of C2H2 class of zinc finger proteins (ZFP)

<220>  
<221> MOD\_RES  
<222> (2)..(3)  
<223> Xaa = any amino acid

<220>  
<221> MOD\_RES  
<222> (4)..(5)  
<223> Xaa = any amino acid, may be present or absent

<220>  
<221> MOD\_RES  
<222> (7)..(18)  
<223> Xaa = any amino acid

<220>  
<221> MOD\_RES  
<222> (20)..(22)  
<223> Xaa = any amino acid

<220>  
<221> MOD\_RES  
<222> (23)..(24)  
<223> Xaa = any amino acid, may be present or absent

<400> 1  
Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa His  
20 25

<210> 2  
 <211> 10  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: ZFP target site  
 with two overlapping D-able subsites

<220>  
 <221> modified\_base  
 <222> (1)..(2)  
 <223> n = g,a,c or t

<220>  
 <221> modified\_base  
 <222> (5)  
 <223> n = g,a,c or t

<220>  
 <221> modified\_base  
 <222> (8)  
 <223> n = g,a,c or t

<220>  
 <221> modified\_base  
 <222> (9)  
 <223> n = a,c or t; if g, then position 10 cannot be g  
 or t

<220>  
 <221> modified\_base  
 <222> (10)  
 <223> n = a or c; if g or t, then position 9 cannot be g

<400> 2  
 nngkngknnn

10

<210> 3  
 <211> 10  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: ZFP target site  
 with three overlapping D-able subsites

<220>  
 <221> modified\_base  
 <222> (1)..(2)  
 <223> n = g,a,c or t

<220>  
 <221> modified\_base  
 <222> (5)  
 <223> n = g,a,c or t

<220>  
<221> modified\_base  
<222> (8)  
<223> n = g,a,c or t

<400> 3  
nngkngkngk

10

<210> 4  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 4  
Asp Gly Gly Gly Ser  
1 5

<210> 5  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 5  
Thr Gly Glu Lys Pro  
1 5

<210> 6  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 6  
Leu Arg Gln Lys Asp Gly Glu Arg Pro  
1 5

<210> 7  
<211> 4  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 7  
Gly Gly Arg Arg  
1

<210> 8  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 8  
Gly Gly Gly Gly Ser  
1 5

<210> 9  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 9  
Gly Gly Arg Arg Gly Gly Ser  
1 5

<210> 10  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 10  
Leu Arg Gln Arg Asp Gly Glu Arg Pro  
1 5

<210> 11  
<211> 12  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 11  
Leu Arg Gln Lys Asp Gly Gly Ser Glu Arg Pro  
1 5 10

<210> 12  
<211> 16  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:linker

<400> 12  
 Leu Arg Gln Lys Asp Gly Gly Ser Gly Gly Ser Glu Arg Pro  
 1 5 10 15

<210> 13  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:ZFP target site  
 region surrounding initiation site of vascular  
 endothelial growth factor (VEGF) gene containing  
 two 9-base pair target sites

<220>  
 <221> protein\_bind  
 <222> (4)..(12)  
 <223> upstream 9-base pair ZFP VEGF1 target site

<220>  
 <221> protein\_bind  
 <222> (14)..(22)  
 <223> downstream 9-base pair ZFP VEGF3a target site

<400> 13  
 agcggggagg atcgcggagg cttgg

25

<210> 14  
 <211> 298  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:VEGF1 ZFP  
 construct targeting upstream 9-base pair target  
 site in VEGF promoter

<220>  
 <221> CDS  
 <222> (2)..(298)  
 <223> VEGF1

<400> 14  
 g gta ccc ata cct ggc aag aag cag cac atc tgc cac atc cag ggc 49  
 Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly  
 1 5 10 15

tgt ggt aaa gtt tac ggc aca acc tca aat ctg cgt cgt cac ctg cgc 97  
 Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg  
 20 25 30

tgg cac acc ggc gag agg cct ttc atg tgt acc tgg tcc tac tgt ggt 145  
 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly  
 35 40 45

aaa cgc ttc acc cgt tcg tca aac ctg cag cgt cac aag cgt acc cac 193  
 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His  
 50 55 60

acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg	241		
Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met			
65	70	75	80
cgt agt gac cac ctg tcc cgt cac atc aag acc cac cag aat aag aag	289		
Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys			
85	90	95	
ggt gga tcc	298		
Gly Gly Ser			

<210> 15  
 <211> 99  
 <212> PRT  
 <213> Artificial Sequence

<400> 15			
Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly			
1	5	10	15

Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg			
20	25	30	

Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly			
35	40	45	

Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His			
50	55	60	

Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met			
65	70	75	80

Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys			
85	90	95	

Gly Gly Ser

<210> 16  
 <211> 298  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:VEGF3a ZFP  
 construct targeting downstream 9-base pair target  
 site in VEGF promoter

<220>  
 <221> CDS  
 <222> (2)..(298)  
 <223> VEGF3a

<400> 16	49		
g gta ccc ata cct ggc aag aag aag cag cac atc tgc cac atc cag ggc			
Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly			
1	5	10	15

tgt ggt aaa gtt tac ggc cag tcc tcc gac ctg cag cgt cac ctg cgc 97  
 Cys Gly Lys Val Tyr Gly Gln Ser Ser Asp Leu Gln Arg His Leu Arg  
 20 25 30  
 tgg cac acc ggc gag agg cct ttc atg acc tgg tcc tac tgt ggt 145  
 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly  
 35 40 45  
 aaa cgc ttc acc cgt tcg tca aac cta cag agg cac aag cgt aca cac 193  
 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His  
 50 55 60  
 acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg 241  
 Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met  
 65 70 75 80  
 cga agt gac gag ctg tca cga cat atc aag acc cac cag aac aag aag 289  
 Arg Ser Asp Glu Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys  
 85 90 95  
 298

ggt gga tcc  
 Gly Gly Ser

<210> 17  
 <211> 99  
 <212> PRT  
 <213> Artificial Sequence

<400> 17  
 Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly 15  
 1 5 10  
 Cys Gly Lys Val Tyr Gly Gln Ser Ser Asp Leu Gln Arg His Leu Arg  
 20 25 30  
 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly  
 35 40 45  
 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His  
 50 55 60  
 Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met  
 65 70 75 80  
 Arg Ser Asp Glu Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys  
 85 90 95

Gly Gly Ser

<210> 18  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: VEGF DNA target  
 site 1 recognition (top) strand

102020 \* 4422268360

<220>  
 <221> protein\_bind  
 <222> (11)..(19)  
 <223> VEGF DNA ZFP target site 1

<400> 18  
 catgcatagc ggggaggagtc gccatcgat

29

<210> 19  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:VEGF DNA site 1  
 complementary (bottom) strand

<400> 19  
 atcgatggcg atcctcccg ctatgcata

29

<210> 20  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:VEGF DNA  
 target site 3 recognition (top) strand

<220>  
 <221> protein\_bind  
 <222> (11)..(19)  
 <223> VEGF DNA ZFP target site 3

<400> 20  
 catgcatatc gcggaggctt ggcatacgat

29

<210> 21  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:VEGF DNA target  
 site 3 complementary (bottom) strand

<400> 21  
 atcgatgcca agctccgcg atatgcata

29

<210> 22  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:primer SPE7

<400> 22  
qagcagaatt cgccaagaag aagcagcac

29

<210> 23  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220> Description of Artificial Sequence:primer SPEamp12  
<223>

<400> 23  
gtggtctaga cagctcgta cttcgc

26

<210> 24  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220> **Description of Artificial Sequence:primer SPE**  
<223> amp13

<400> 24  
ggagccagg ctgtggtaaa gtttacgg

28

<210> 25  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer SPEamp11

<400> 25  
ggagaagctt ggatcctcat tatccc

26

<210> 26  
<211> 83  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:sequence  
ligated between XbaI and StyI sites

<400> 26 tctagacaca tcaaaaaccca ccagaacaag aaagacggcg gtggcagcgg caaaaagaaa 60  
cagcacatat gtcacatcca agg 83

<210> 27  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:primer GB19

<400> 27  
 gccatgccgg tacccatacc tggcaagaag aagcagcac 39

<210> 28  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:primer GB10

<400> 28  
 cagatcggat ccacccttct tattctggtg ggt 33

<210> 29  
 <211> 589  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:designed  
 6-finger ZFP VEGF3a/1 from KpnI to BamHI

<220>  
 <221> CDS  
 <222> (2)..(589)  
 <223> VEGF3a/1

<400> 29  
 g gta ccc ata cct ggc aag aag cag cac atc tgc cac atc cag ggc 49  
 Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly  
 1 5 10 15

tgt ggt aaa gtt tac ggc cag tcc gac ctg cag cgt cac ctg cgc 97  
 Cys Gly Lys Val Tyr Gly Gln Ser Ser Asp Leu Gln Arg His Leu Arg  
 20 25 30

tgg cac acc ggc gag agg cct ttc atg tgt acc tgg tcc tac tgt ggt 145  
 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly  
 35 40 45

aaa cgc ttc aca cgt tcg tca aac cta cag agg cac aag cgt aca cac 193  
 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His  
 50 55 60

aca ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg 241  
 Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met  
 65 70 75 80

cga agt gac gag ctg tct aga cac atc aaa acc cac cag aac aag aaa 289  
 Arg Ser Asp Glu Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys  
 85 90 95

gac ggc ggt ggc agc ggc aaa aag aaa cag cac ata tgt cac atc caa 337  
 Asp Gly Gly Ser Gly Lys Lys Lys Gln His Ile Cys His Ile Gln  
 100 105 110

ggc tgt ggt aaa gtt tac ggc aca acc tca aat ctg cgt cgt cac ctg	385
Gly Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu	
115 120 125	
cgc tgg cac acc ggc gag agg cct ttc atg tgg acc tgg tcc tac tgt	433
Arg Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys	
130 135 140	
ggt aaa cgc ttc acc cgt tcg tca aac ctg cag cgt cac aag cgt acc	481
Gly Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr	
145 150 155 160	
cac acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc	529
His Thr Gly Glu Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe	
165 170 175	
atg cgt agt gac cac ctg tcc cgt cac atc aag acc cac cag aat aag	577
Met Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys	
180 185 190	
aag ggt gga tcc	589
Lys Gly Gly Ser	
195	

82163501  
 <210> 30  
 <211> 196  
 <212> PRT  
 <213> Artificial Sequence

<400> 30	
Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly	
1 5 10 15	
Cys Gly Lys Val Tyr Gly Gln Ser Ser Asp Leu Gln Arg His Leu Arg	
20 25 30	
Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly	
35 40 45	
Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His	
50 55 60	
Thr Gly Glu Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met	
65 70 75 80	
Arg Ser Asp Glu Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys	
85 90 95	
Asp Gly Gly Ser Gly Lys Lys Gln His Ile Cys His Ile Gln	
100 105 110	
Gly Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu	
115 120 125	
Arg Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys	
130 135 140	
Gly Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr	
145 150 155 160	

His Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe  
 165 170 175

Met Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys  
 180 185 190

Lys Gly Gly Ser  
 195

<210> 31  
 <211> 42  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:JVF9 VEGF3a/1  
 target oligonucleotide

<400> 31  
 agcgagcggg gaggatcgcg gaggcttggg gcagccgggt ag

42

<210> 32  
 <211> 42  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:JVF10 VEGF3a/1  
 target oligonucleotide complementary sequence

<400> 32  
 cgctctaccc ggctgccccca agcctccgcg atcctcccccg ct

42

<210> 33  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:primer JVF24

<400> 33  
 cgcggatccg cccccccgac cgatg

25

<210> 34  
 <211> 62  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:downstream  
 primer JVF25

<400> 34  
 ccgcaagctt acttgtcatac gtcgtccttg tagtcgctgc cccccaccgta ctcgtcaatt 60

cc

62

<210> 35

<211> 7

<212> PRT

<213> Simian virus 40

<220>

<221> PEPTIDE

<222> (1)..(7)

<223> SV40 large T antigen nuclear localization sequence  
(NLS)

<400> 35

Pro Lys Lys Lys Arg Lys Val

1

5

<210> 36

<211> 61

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:segment from  
EcoRI to KpnI containing Kozak sequence including  
initiation codon and SV40 NLS

<400> 36

gaattcgcta gcgccaccat ggcccccaga aagaagagga aggtggaaat ccatggggta 60

c

61

<210> 37

<211> 187

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:segment from  
KpnI to XhoI containing BamHI site, KRAB-A box  
from KOX1, FLAG epitope and HindIII site

<400> 37

ggtaccggg gatccggac actggtgacc ttcaaggatg tatttggta cttcaccagg 60

gaggagtgga agctgctgga cactgctcag cagatcgatg acagaaatgt gatgctggag 120

aactataaga acctggttc cttgggcagc gactacaagg acgacgatga caagtaagct 180

tctcgag

187

<210> 38

<211> 277

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:inserted  
fragment from BamHI to HindIII sites

<400> 38  
 ggatccgccc ccccgaccga tgcgtcgctg ggggacgagc tccacttaga cggcgaggac 60  
 gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 120  
 gattccccgg ggccgggatt taccggccac gactccgccc cctacggcgc tctggatatg 180  
 gccgacttcg agttttagca gatgtttacc gatgcccttg gaattgacga gtacgggtggg 240  
 ggcagcgact acaaggacga cgatgacaag taagctt 277

<210> 39  
 <211> 118  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Description of Artificial Sequence:sequence  
 replacing NLS-KRAB-FLAG with NLS-FLAG only  
 <400> 39  
 gaattcgcta gcgccaccat ggccccaag aagaagagga aggtgggaat ccatgggta 60  
 cccggggatg gatccggcag cgactacaag gacgacgatg acaagtaagc ttctcgag 118

<210> 40  
 <211> 204  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Description of Artificial Sequence:insert into  
 MluI/BglII sites of pGL3-Control to create  
 pVFR1-4x  
 <400> 40  
 acgcgttaagc ttgttagcga gcggggagga tcgcggaggc ttggggcagc cggtagagc 60  
 gagcggggag gatcgccggag gcttggggca gccgggtaga gcgagcgggg aggatcgccg 120  
 aggcttgggg cagccggta gagcgagcgg ggaggatcgc ggaggcttgg ggcagccggg 180  
 tagagcgctc agaagcttag atct 204

SCANNED, #14